
AI GENERATED CONTENT AND THE LAW OF DEFAMATION IN INDIA

Manya Singh, Symbiosis Law School, Noida

ABSTRACT

Artificial intelligence's rapid rise has impacted the production of digital content, but it has also brought up significant legal issues, especially in relation to defamation. Artificial intelligence (AI) tools, such as deepfake technology, can produce incredibly lifelike but fraudulent media, frequently causing harm to reputation. This article examines the shortcomings and ways in which Indian defamation law, in particular Section 356 of the Bharatiya Nyaya Sanhita, 2023, addresses such harm.

With an emphasis on Section 356 of the Bharatiya Nyaya Sanhita, 2023, this article explores the legal issues raised by AI-generated content in the framework of Indian defamation law. Current legal frameworks find it difficult to assign liability for defamatory content produced by AI because it lacks legal personhood and cannot possess intent (*mens rea*). It illustrates the practical consequences of AI-driven defamation and the pressing need for regulatory clarity by analysing recent cases like *Arijit Singh v. Codible Ventures LLP* and *Aaradhya Bachchan v. Bollywood Time*.

Despite the gravity of the problem, India does not yet have a specific legal framework to control misinformation produced by AI or its effects. To lessen the dangers posed by synthetic media, the paper makes the case for swift legislative changes, more stringent content regulations, and public awareness initiatives. In an era driven by artificial intelligence, it also highlights the significance of global collaboration, platform responsibility, and the legal acknowledgement of digital harm.

Keywords: Artificial Intelligence, Defamation, Bhartiya Nyaya Sanhita, Deepfake, Legal personhood.

Introduction

The term artificial generated content refers to the use of generative AI algorithms to create content such as text, image, video etc. these AI models offer massive amounts of content as per the needs of the user with a low production cost and in an efficient manner. The concept of AI-generated content is relatively new for the people and is being widely misused due to lack of legal mechanism present for the people to seek relief in case they are being made the victim of the content generated using artificial intelligence, we shall discuss more of this in the analysis part of this research project.

Defamation is a public communication which tends to injure the reputation of another. Defamation is when someone makes false statements about another person that harm their reputation. Tort of defamation can be considered as a man's inherent personal right which every man has a right to preserve.¹

In the current generation where Artificial intelligence tools have renovated content creation, enabling prompt creation of images, texts, videos and more. There is no doubt in how efficient and cost-effective these tools are but there is a significant question pertaining to the possibility of distorted information which may even lead to serious harm to reputation, defamation. Maneuvering this new environment calls for an understanding of the ethical and legal consequences of defamation and ai generated content. This assignment aims to analyze the relation between defamation and ai generated content, it will concentrate on how there aren't any legal protections available against this content and how it can be defamatory.

What is Defamation

Defamation is mentioned under section 356 of the Bharatiya Nyaya Sanhita, 2023.. This section replaces the old section 499 of the Indian penal code.

As per this section defamation means publishing any statement, either spoken or written, which intends to harm the reputation of the person it is directed towards.²

For a statement to be defamatory, it must fulfill following essentials:-

¹ *Ram Jethmalani v. Subramaniam Swamy*, AIR 2006 DEL 300.

² Bharatiya Nyaya Sanhita, § 356, No. 45, Acts of Parliament, 2023 (India).

- the statement must be false
- it must be directed towards the plaintiff
- it must hurt or tarnish their reputation
- the statement must be published

However, there are a few cases where the statement, even though defamatory in its ordinary sense, would not be considered defamatory. The exceptions are as follows:-

- Truth- if what you said is true, then one cannot be held liable under defamation.
- Fair opinion- if you are giving your honest opinion on something or someone, then it won't constitute defamation. However, the opinion must be made in good faith and should be based on the existing facts and not on hypothetical or untrue facts.
- Criticism of conduct of a public servant- it is not defamation if you honestly express your opinion on how someone acted in a public matter, however, the statement must be on their character based on that behaviour and not beyond that.
- Report of court cases- if you publish the proceedings of the court accurately, it won't amount to defamation.
- Merits of a case- expressing opinion, in good faith, about the merits of the case already decided by the court is not defamation.
- Merits of public performance- Expressing opinion, in good faith, on the merits of a public performance like literary or artistic work, is not defamation.
- Censure by authority- it is not defamation when a person under an authority, by law or contract, censures someone under their authority for a misconduct. Provided that such censure is made in good faith and is within their professional relationship.
- Accusation- Accusation made in good faith, to protect one's own interest, to an authorised person is not defamation.

-Imputation- Imputation made for the protection of one's own interest or someone else's is not defamation as long as it is made under good faith.

-Caution- Statements made under good faith as a caution against potential harm for safety or precautionary purposes are not defamation.

Now that we know what defamation is in the legal sense, let us now understand the concept of deepfake and AI generated content and the risk it possesses.

Concept of AI generated content and deepfake

As we have briefly discussed about the AI technology in the introduction of the paper, let us understand another aspect of it that is deepfake technology. It is a type of AI based technology that uses machine learning algorithms, particularly generative adversarial network (GANs).³ The goal of this type of technology is to create highly realistic synthetic media that resembles real people with manipulated content.¹ This realism offered by this technology could be harmful for people especially for those who have a significant social media presence, and their images and videos are available easily. The core problem arises in the absence of a legal framework on which people could rely to seek justice and protect their rights.

Risk through AI

AI-generated content, particularly through technologies like deepfakes, poses serious risks of defamation and identity theft. These technologies can be used to create false representations of individuals, allowing their identity to be stolen and manipulated. Through deepfakes, forged images or videos can be made to depict a person doing or saying things they never actually did. This technology enables the spread of misinformation, and such manipulated content can be used to sway public opinion, create false narratives, and severely damage a person's reputation or goodwill. Furthermore, deepfake technology makes it easier to add voice overs or alter videos, making it possible for people to spread harmful, defamatory content, either targeting a specific individual or spreading hatred about others. In some cases, this technology is used maliciously for serious crimes, such as child pornography or political propaganda, exacerbating the harm caused by the fabricated content. The damage inflicted through these false

³ Yuntao Wang, *A Survey on ChatGpt: AI-Generated Contents, Challenges, and Solutions*, 4 *IEEE Trans. Artif. Intell.* (2023)

representations can have a long-lasting impact on an individual's public image, and without a robust legal framework to address the misuse of such technologies, individuals are left vulnerable to significant reputational harm.

AI has no personhood

In today's day and age where artificial intelligence is capable of generating image or video as per the description given as the input by the user, it does not possess legal personhood, necessary for it to be subject to any rights and duties. For someone to be held liable for any offence under the Bhartiya Nyay Sanhita or the erstwhile Indian Penal Code, it is important to establish that the offender had both, the guilty mind(*mens rea*) and the guilty intention(*actus rea*). However, the AI lacks consciousness, awareness and intention and hence cannot be said to possess *mens rea* like the natural persons do. Furthermore, section 3(42) of the general clauses act, 1897 defines persons and includes all the natural persons and legal entities like the companies but has no mention of the machines or the new AI systems.

The grey zone and how

Now a question may come to your mind, that as we know that AI lacks consciousness and it can't generate an image or a video without human intervention or input, then if someone posts a content generated through, defamatory in nature, and posts it online, then wouldn't it be easy to trace them through their IP address? Yes, in theory, it will be easy to trace them but here is the catch, it is not always that easy. It can be tough to trace people's IP address when:-

- they have used tools like VPN or proxy address to hide their real IP address.

- they post it from a public place, like a cafe

This creates a grey area both socially and legally. With rapid rise in the AI generated content, a person unfamiliar with this aspect may genuinely believe that the person in the video or image is actually saying or doing thing that is defamatory in nature, when in reality the video is entirely fabricated and that person has no relation to their acts as shown in the content so generated. The damage however is done within minutes, even before the real truth comes out. It is clear that the impact of defamatory content through AI is real and only growing, however there is still no dedicated legislation governing this grey aspect, let us see the laws that are currently prevailing in India to address this issue.

The Law on AI

Artificial intelligence and its use has seen a rapid growth in the Indian market, but despite all this there is no proper legislation governing its market or its legal implications. While the BS or the IT Act govern the legal issues relating to it, even then there is no specific mention of AI in it. This gap in the law creates uncertainty and raises various questions regarding accountability, responsibility and regulation- especially when AI is all potent to create content using someone's voice or image without that person's consent. However there have been a few judicial pronouncements on these lines.

Arijit Singh v. Codible Ventures LLP and Others

In this case, Arijit Singh, the plaintiff, a renowned singer who has built immense goodwill and reputation as a playback singer who is being subject to exploitation by the defendants using AI-generated content. AI tools are used to synthesize the plaintiff's voice, mimicking his sound and mannerisms, while deepfake and face morphing technologies are used for commercial gain. Additionally, manipulated GIFs from his videos have subjected him to ridicule, harming his reputation. In this case, the court held that the plaintiff's name, likeness, image, persona, etc., deserves to be protected not only for his own sake but also for the sake of his family and friends who would not like to see his image, name and other elements being misused, especially for such tarnishing and negative use. Court was of the view that defendants are attracting visitors to their websites and AI platforms by capitalizing on plaintiff's popularity and reputation, thereby subjecting the plaintiff personality rights to potential abuse, counterfeit sound recordings and videos that misuse plaintiff's character and identity, creation of audio/video using AI without plaintiff's consent and commercially using the same could potentially jeopardise plaintiff's career.⁴

Aaradhya Bachchan and Another v. Bollywood Time and Another

In this case the defendant posted a video on famous social media platform YouTube stating that the plaintiff, a 11-year-old child of famous Bollywood actors Abhishek Bachchan and Aishwarya Rai, is no more for mere publicity, even morphed images to support the misinformation. Parents of the plaintiff prayed for the temporary injunction from uploading

⁴ *Arijit Singh v Codible Ventures LLP and Others*, [2024] BOM HC

any videos through social media platforms which result in tarnishing and bringing disrepute to the goodwill and reputation that is vested in the name of Bachchan family and defamation of the plaintiffs.⁵

Above cited cases clearly present before us points which show the involvement of AI-generated content posing a considerable harm to the respective plaintiff's reputation and goodwill, these points are as follows:

- Involvement of AI in generating content which exposes plaintiff to hatred
- Poses injury to trade or profession of plaintiff (as could be seen in the case law of Arijit Singh and Gaurav Bhatia)
- The content so generated shows the plaintiff doing such things which anyone in the first glance would think that the plaintiff is actually doing.
- The content so generated could be defamatory in nature if it shows the plaintiff doing certain things which are actually demeaning in real life and decreases the reputation of the plaintiff in the mind of right-minded people in the society.

Suggestions:

The current legal framework to tackle such offences is not adequate in India and is not fully discussed. The provisions pertaining to such technology should be added if it acts to regulate this deepfake technology effectively. There is need to update laws in this domain and as well as there is urgent need to address the issue of defamation spread through deepfake. The problem is not just limited to India but is an international one and international cooperation would be needed. It also becomes extremely important for the government to educate people about potential risks attached to such technology and how it is dangerous for any individual's reputation. The government should:

- a. Block public access to such misinformation, this will help in controlling misinformation from spreading to many people.

⁵ *Aaradhya Bachchan and Another v. Bollywood Time and Another*, [2023] DEL HC

- b. punitive approach for reputational harm caused by such technology especially if a child is the victim of such an act
- c. imposing responsibility on intermediaries.

Conclusion

The rapid growth of artificial intelligence in the field of content creation has offered numerous benefits like enabling the production of text, images and videos with remarkable efficiency. This has also introduced serious ethical and legal concerns pertaining to defamation. AI-generated content has the potential to manipulate reality creating a false representation of individuals that can be seriously misleading and can hamper the reputation. Defamation, a long-established legal principle, safeguards individuals from false statements harming their reputation. AI-generated content often meets the criteria for qualifying as defamatory, but the challenge lies in determining the accountability, whether the accountability rests with AI-developers, the users who generate the content, or the platforms that distribute it. The lack of provision regarding AI-generated defamation makes it difficult for victims to seek justice. The cases discussed the analysis part of this project underscore the urgent need for legal reforms to address the misuse of AI. Currently, Indian Law does not adequately regulate the AI-generated content leaving a significant gap in legal protection. Government should take proactive measures such as blocking access to misinformation, imposing stringent penalties for harm caused to reputation using AI and holding intermediaries responsible for hosting such content on their platform. Hence, immediate action is needed to establish a legal framework that effectively addresses the challenges posed by AI-generated defamation.